Block I: Chapter I:

Testing and assessment will help the teacher whether the learning objectives have been achieved or not.

Learning Experiences.

Evaluation procedures: whether I am using a quiz, or mcq or descriptive answers.

Evaluation checks whether the teaching materials appropriate? whether the approach of method is effective, it gives information about the feedback. Evaluation helps us understand whether learning happened or not and helps us find the reasons for the failure. Evaluation is a continuous process much like learning. When evaluating we need to measure all the members of a class with the same yardstick: the same test paper, the same time, the same conditions. The testing tells us how successful the instruction has been and what modifications are necessary. This is called the washback effect of testing on teaching.

We need to test because we need to understand how much learning is happening. We need to see how a group of students are able to perform. How well they are able to follow instructions, or whatever is being taught to them.

The traditional evaluation methods focused more on the content reproduction. Mostly memory-based questions.

What is a language Testing: the stages in construction. A particular test is often referred as an instrument. In language testing we are dealing with the ability to use language correctly and appropriately. Language ability used to be tested indirectly from the responses on the content based questions on essays, novels, short stories etc. In recent years language testing has shifted from 'content' to skills. The focus of a language should be functional use of language and not the aspects of language. Language test should include items that cover different areas of language ability: Listening, speaking, reading and writing and through them the integral elements of grammar and vocabulary.

What are the stages of test construction.

- Specify the skills and abilities to be tested.
- Describe the language through activities. (Speaking skills activities.)
- Produce the test and administer it under specified conditions. (Time, limitations, things you have to outline before)

- Assess the quality of the responses of the test-takers. (How well students have performed)
- Express these assessments on a scale, in terms of grades, or percentage.
- Interpret the final result.

Note: The question paper is called a test instrument in technical term of testing.

Stages of language test construction:

Specify the skills and abilities to be tested Decide exactly what you want to measure (e.g., listening comprehension, grammatical accuracy, speaking fluency). This stage ensures the test targets the intended construct of language proficiency.

Specify the skills and abilities to be tested.

- define the scope language performance,
- the particular levels of the learner's abilities,
- the context of the target language use,
- the purpose for which the learners have to use language,
- the genres that will serve these purposes. (whether a short story is suitable for producing that language skill or a poem or a novel or a conversation. So genre that will serve the purpose has to be identified.

Test: Placement test for adult immigrants starting college ESL classes

- **Levels**: Beginner to Intermediate (CEFR A2–B2)
- **Skills tested**: Listening, Reading, Writing, Speaking (the particular level of Skill)
- **Where they will use English**: College classes + daily life (doctor, bank, shopping, registration) (the context of language use)
- **Why they need English**: Attend lessons, do homework, talk to teachers, manage everyday tasks (The purpose for using the language.
- **Text types used in the test**: (the genre that will serve the purpose)
    - Listening: short talks, announcements, phone messages
    - Reading: signs, emails, forms, short articles

- Writing: short email or paragraph (120–150 words)
- Speaking: describe a picture, answer questions, short role-play

Describe the language through activities that involve the selected abilities Define the content and contexts (topics, text types, situations) that will elicit the target skills. This creates a clear blueprint linking abilities to real-language use.

Devise efficient test tasks Design specific item types and prompts (multiple-choice, gap-fill, oral interview, essay, etc.) that effectively measure the abilities. Tasks must be practical, clear, and appropriate in difficulty and time.

Produce the test Write, edit, and format all items, instructions, answer sheets, and rubrics into a final test version. This stage includes piloting or pre-testing is often included to check quality.

Administer it under specified conditions Give the test to candidates following strict, standardised procedures (time limits, silence, no aids, etc.). Consistent administration ensures fairness and comparability of results.

Assess the quality of the responses of the test takers Score the answers using marking keys or rating scales; for productive skills, trained raters apply analytic or holistic rubrics. This stage may include double-marking and moderation for reliability.

Express the assessment on a scale and interpret the final result Convert raw scores into grades, percentages, band levels (e.g., CEFR A1–C2), or pass/fail decisions. Provide meaningful interpretation of what the score says about the test taker's language ability.

Testing validity is the extent to which a test accurately measures what it intends to measure. E.g. if I wanted to test speaking skills then the task and activities I design should elicit speaking skills, then only it will have testing validity. After identifying the skill, I need to describe the skill and sub-skill. Th

When forming a question paper (test instrument) following things have to be taken into consideration.

You need to have a proper layout. You need give clear instructions, give clear instructions in the questions. there are two types of instructions general ones and instructions specific to questions. Testing

administrative conditions should be good. Proper seating arrangement, hand out the questions paper collect the answer sheets, take the attendance. Sometimes specific time is given to specific section or questions. There must be inter- rater reliability. The answer key has to be very corrected. The criteria should be applied to all the responses. The scoring key is to be given.

Purposes of testing and kinds of tests:

Purpose: to measure language proficiency, to discover how successful learners have been in achieving the objectives of a course of study. To diagnose the learners' strengths and weaknesses, to identify what they know and what they do not know. To assist placement of students by identifying the stage.

Proficiency test: It is nothing but language proficiency test. It tests the ability to use language effectively in the real world. It implies balanced development of the four LSRW skills, grammar and vocabulary. There are usually three levels of language proficiency. Tests like IELTS, TOFEL are type of proficiency tests. The content of a proficiency test are reading comprehension, writing sections, grammar, vocabulary etc. Proficiency tests are not designed based on any syllabus or course of instruction. Below is the level of language proficiency.

1) Basic user level
2) Independent user level
3) Proficiency level (C1 and C1)

Achievement test: What has the learner been able to achieve. Achieve tests are based on syllabus. We need to keep in mind that the syllabus has teaching objectives and the materials for the realization of the objectives. The aim of the teacher is to know whether the learners have learnt what was taught. Formative and summative assessment are part of this achievement tests. Class tests and classroom-based evaluation are ongoing and give you insights into the learning process and help you modify your instructions according to the feedback you receive.

Diagnostic Test: Designed to understand whether someone has learnt something or not. The tests identifies whether a group of students has acquired a particular aspect of language or not. The purpose of a diagnostic test is to identify the strengths and weaknesses of a learner, so that feedback can be utilized for ensuring effective learning.

Placement Test: Locate the level of proficiency. It helps us determine the levels of learners and put them in two or three groups facilitate learning.

Proficiency Test	Measures language proficiency, fluency and

	accuracy
Achievement Test	To discover how successful students have been in achieving the objectives of a course of study
Placement test	To assist placement of students by identifying the stage or part of a teaching programme
Progress	To see the progress that learners have made with learning.
Diagnostic test	To diagnose students' strengths and weaknesses, by identifying what they know and what they do not know.

Objectives of Teaching and Testing: Chapter II. Block I:

In the beginning of each block there is something called after studying this unit you will be able to learn so and so things. This is called instructional goal. What is being taught.

The instructional goals should clarify three things.

- What the learners have to do? (Learner activity)
- What are the conditions in which the activity will be done (learning environment)
- Wha is the quality of the performance expected of the student (learning outcome)

Bloom's taxonomy's educational objectives serve as the accepted model for specifying objectives for various levels of instruction (teaching) the levels of blooms taxonomy for education. These stages are cognitive levels. Remembering is the first stage and creating is the highest last stage of cognitive level involved in learning. The final aim of a teaching instruction is to make the learners create something new based on what is taught to them.

Creating (Highest Stage)	*Generating new ideas, products or ways of views things, finding new perspectives, viewing things, designing, constructing, planning, producing and inventing*.
Evaluating	*Justifying a decision or course of action, checking hypothesizing, critiquing, experimenting, judging assessing,*
Analysing	*Breaking information into parts to explore, understandings and relationships, differentiate, comparing information, organizing, deconstructing, interrogating, finding etc*.
Applying	*Using information in another familiar situation*, *implementing, carrying out a task, using, and executing information. Information transfer, create or produce*.

Understanding	*Explaining ideas or concepts, interpreting, summarising, paraphrasing, classifying,*
Remembering (Lowest stage)	*Recalling information, recognizing, listing, describing, retrieving, naming, finding.*

Based on the cognitive levels in Blooms taxonomy we need to look at linguistic objectives in a language syllabus and language instruction. In language syllabus linguistic and cognitive objectives are integrated.

J.B. Carol's classification of different areas of language provides us with a general framework for deriving objectives. Carol's table for linguistic competences.

Mode/level	Listening (receptive passive skill)	Speaking (Active, productive skill)	Reading (receptive passive skill)	Writing (Active, productive skill)
Phonology Pronunciation and spellings	Sound discrimination, stress and intonation, recognition through listening tasks. Minimal pair, listening to rhymes, mcq	Pronunciation practice through drills, dialogues and role plays. Test tasks:	Sound-symbol relationship, recognition, reading aloud.	Spelling dictation, transcription activities
Grammar	*Identifying grammatical structures in spoken texts*	Using grammar accurately in speech through role plays and interviews.	Recognizing grammatical forms while reading.	Producing grammatically sentences and paragraphs.
Lexicon/ Vocabulary	Understanding word meanings through contextual listening.	Using appropriate vocabulary in spoken communication.	Inferring meaning of words from reading texts.	Using vocabulary effectives in writing tasks
Integrative skills (LSRW, grammar and Vocabulary integrated.	Listening comprehension integrating phonology, grammar and vocabulary	Fluent speaking integrating all language components.	Reading comprehension requiring integrated language knowledg	Coherent integrating grammar, vocabulary organization.

			e.	

Listening, reading, speaking and writing are the modes for testing of one of the levels. (Phonology, Grammar, and vocabulary. Sometimes two or more (LSRW) skills may be involved in the testing of each element.

The objectives of listening skills involve:

- The students can take down dictation of simple sentences (phonology)
- Can retell the contents of authentic audio texts.
- Can understand the main line of a story listened to.
- Can follow simple instructions.
- Can take down important points from a talk/lecture.

Element	Mode	Objective
Vocabulary/ Lexis	Reading/Writing	Is able to identify collocations
Grammar	Reading	Is able to identify errors in subject-verb agreement
Phonology	Listening	Is able to discriminate by similar sounding words
Vocabulary	Reading	Is able to understand the meanings of idiomatic expressions.

Both speaking and listening are inter-related activities. Speaking is usually a response to listening in interpersonal communication. At the basic level of speaking ability pronunciation of the distinctive sounds and sound patterns of the second language have to be acquired. The teaching and testing of speaking can follow the list of objectives given below.

- Can produced correct stress and tone patterns.
- Can respond clearly at normal conversational speed.
- Can respond immediately when spoken to.
- Can speak coherently on general topics using appropriate linkers
- Can use vocabulary suitable to the situation or context.
- Can speak with a reasonable degree of grammatical accuracy.
- Can give directions to specific places on the telephone.
- Can participate in role plays and speak at normal conversational speed
- Can *read out sentences with appropriate stress and intonation*
- Can respond to questions played on the radio

There are four levels of reading comprehension: global, Factual, lexical and inferential.

Reading comprehension for lexical and factual comprehension relate to the cognitive levels remembering and understanding. Global comprehension relates to synthesizing. Inferential reading relates to analyzing and evaluating. The cognitive level for reading is mostly understanding as reading fosters comprehension.

The aims of reading skills: The purpose of reading.

- Reading for decoding sound and the spelling.
- Reading for meanings -knowing meanings of words, grammatical forms, sentences structures and cohesive linkers.
- When we read for factual and specific information it is scanning.
- When we read for overall understanding it is skimming.
- When we read intensively to understand theoretical concepts it is intensive reading
- When we read for pleasure such as novels, poems etc. it is extensive reading.
- Reading can be combined with writing, making notes, summarising, transfer of information activities.

Purposes for Reading	Levels of Reading Comprehension
1. Overall understanding 2. Understanding factual details 3. Understanding underlying meaning 4. Understanding word meaning	a. Lexical comprehension b. Inferential comprehension c. Actual comprehension d. Global comprehension e. Intensive comprehension

Writing: It is a productive skill. Usually numbers of errors in spelling, punctuation and grammar are the criterion. Some of the common tests are: Writing is tested on the basis of clarity, coherence, precision of thought, development of ideas, and vocabulary use.

- Make sentences using the given words
- Rewrite this passage in the third person.
- Complete the story with suitable words.
- Use the outline to write a story/essay
- Describe the picture

- Write an essay, write a notice, etc. These activities elicit writing skills.

Pedagogic tasks (takes place in classroom while teaching) are those that are used in the classroom during instruction simulating real life communication. They are open-ended and are not limited by time and place. Test tasks, in contrast are time-bound. They have to be completed in the test venue. A test is conducted under strict supervision. We must remember that there is a distinction between teaching tasks (during the course of instruction) and testing tasks (Those used in exams to test language abilities)

Kinds of Testing: Language ability can not be tested directly, it is tested indirectly through performance, and we can only test performance not competence.

Direct- testing: Direct testing measures language ability by requiring the student to perform the actual skill being tested. While we cannot see mental competence, we can directly observe the performance it produces. The evidence of the student's ability is directly available through their performance. (Can they write a letter?) Speaking skills can be directly tested through oral interviews, describing a picture, talking about one's family. Writing skill can be tested through writing a notice or an essay. Pronunciation can be tested directly through reading a passage aloud.

Indirect testing: Indirect testing measures the underlying abilities, sets of knowledge that are assumed to support the main language skills. It is useful for diagnostic information. (checking if a student understands subject-verb agreement) Writing can be tested indirectly through arranging scrambled sentences. Grammar can be indirectly tested through identifying an error in a written sentence.

Discrete-Point Testing: The approach is based on the idea that language can be broken down into individual elements. It aims to assess specific linguistic knowledge. E.g. vocabulary, grammar, phonology in isolation. These tests are often indirect. It has three types.

Context-free tasks: Matching a phrase with its meaning. (matching 'look into' with 'examine'.

Context-embedded tasks: Fill-in-the blank sentences where the answer relies on knowing the specific word meaning rather than the context itself.

Structure-focused: MCQ questions testing a single tense or preposition.

Integrative testing: It aims to assess the learner's ability to use multiple language elements and skills simultaneously. It tests overall language ability. Requires learners to combine grammar, vocabulary, etc with LSRW. These tests are direct, because they mirror real-world language use. Writing an essay based on a provided prompt can be an example of integrative testing.

Norm-Referenced Testing: It is designed to compare a student's performance against a norm group. The primary goal is to rank students and determine their relative standing. It covers a wide range of abilities such as general English fluency. The Percentile and percentage style of marking is norm-referenced. A norm-referenced test identifies whether the test taker performed better or worse than other test takers. The goal of norm-referenced test is to distinguish between the ability levels of the members of a group. It does not describe what a learner can do or cannot do.

Criterion-referenced Test: These are those that describe in clear terms whether or not an individual candidate can or cannot perform a task. The tests indicate what the candidate is able to do. It is something like the cut-off percentage we have for being selected for certain professional courses. **Why criterion-referenced fits better**

It lets you see whether learners have actually mastered specific language skills and objectives (e.g., "can write a coherent paragraph," "can follow everyday conversations"), which aligns directly with communicative language teaching and curricula. It provides clearer diagnostic information and feedback for improvement, skill by skill, rather than just a relative rank against other test takers.

Objective and Subjective Testing:

Objective type test items have only one right answer. The score, the answer will be the same. Objective type of testing reduces subjectivity in the examiners. These type of tests are close-ended. Objective tests are almost always discrete point tests and therefore to a certain extent indirect. Subjective tests are open-ended, when an essay or a paragraph is scored by two or three examiners, there may be a variation in the assessment. Therefore, subjective assessment are tend to be less reliable than objective assessment.

Formative Assessment and Summative assessment:

Formative assessment. Periodic tests are a feature of all school courses. Continual feedback is to be given so that learners perform

better on subsequent tests. In formative assessment the progress that a student makes over a period of time is considered important and is recorded. There is a scope for feedback which is used for improvement.

Summative assessment is the end-of-the term semester year examination. It includes all that has been done in a year. Usually administered by an external body. Achievement tests are summative in nature.

What Makes a good Test: Unit 3:

A good testing is based on certain general principles of testing. A test is an instrument, a measurement and we have to ensure that if the instrument used is reliable and authentic. The two most important terms associated with this truth value of test-information are validity and reliability. Validity, reliability, authenticity, interactiveness, practicality and impact are important principles of language testing.

We test because we wish to know if the learning objectives of a particular area of language ability have been fulfilled or not. It should be observed that if the test being used is reliable or not. Various criteria are available to judge the quality of a test are as follows.

- Is the test valid
- Is the test reliable or not?
- Is the test authentic?
- Is the test interactive?
- Is the test practical?
- Does it have a washback effect?

Please remember that no test is perfect. The goal is to constantly improve the validity, reliability and usability of the tests.

Test validity: Every test has a specific purpose. How well does the test fulfil this purpose? Does it accurately measure what it seeks to measure? The course of the Testing Language and Literature then the questions also should be from the course itself to make the test valid. It should measure what was taught to students. Test validity includes the test design, the scoring procedure, and the interpretation of test performance. There are three kinds of validity.

- Construct validity
- Content validity
- Criterion validity

Construct validity: It asks whether the adopted test measures the psychological test it aims to measure. Construct means the mental

ability. It has something to do with ability, intelligence and mental abilities. Reading, writing, counting spelling, etc are constructs. Construct validity ensures that the test truly reflects the theoretical idea behind it rather than measuring something else. E.g. Suppose the test is designed on speaking skills, then theoretical concepts used to test speaking skills should include fluency, pronunciation, grammar and vocabulary. It is easy to say a test has construct validity in a direct test say of writing or speaking. You can observe it properly directly. If we test reading comprehension, we must be sure that the items test only comprehension and not the grammaticality of the written answers.

Content Validity: It related to the content of the test. Content validity refers to whether the test items are relevant to the topics taught in the course. It determines if a test accurately represents the content it aims to measure, means the questions reflect the material studied. Only that should come in the exam that have been taught in the class. No out of the syllabus thing. The test must cover everything, every content. It should be fair, representative representation of all the topics. The abilities stated in the objectives should form the content of the language test. The items selected should be overall representation of all the topics covered within a course.

Criterion Validity: It assesses how well a new test measures its intended purpose. There should be a correlation between the criteria to interpret the question papers. **criterion validity** is about proving a test works by comparing it to something else that we already trust.

Concurrent Validity: It examines the relationship between a new test and an established test. Concurrent validity means the same construct at the same time.

Predicative Validity: It evaluates how well a test forecasts future outcomes. When the teaching is aimed at developing skills that sustain in future. Ph.D. competitive exams, entrance tests.

Predicative Validity	*The correlation between an individual's performance on a test and his performance on a test on future t*hat are represented in the test.
Face Validity	*The correspondence between the appearance to the test content and the ability it seeks to test.*
Construct	*A mental ability that can be defined and can be shown to exist or not exist in an individual*
Content Validity	*The quality of a test that ensures a fair and representative coverage of the items that can reflect an ability. (If my objective is to teach speaking skills then the content I select as my material should reflect*

	speaking skills.
Concurrent Validity	*The correlation between performance on a shorter representative version of a longer test.*
Construct validity	*The quality of a test that indicates fairly accurately, the existence or non-existence of an ability in an individual.*

Reliability: It refers to the consistency and stability of a measurement. If you give the same test to the same person under the same conditions, they should get a similar result. If we evaluate a student's portfolio, collection of writing it may be reliable. If there are multiple versions of the same content over a period of time, if the results remain the same then the test is reliable.

Scorer Reliability: the person who checks the question paper.

Objective tests usually have high reliability because there is only one right answer. In subjective tests the reliability is lower.

Arthur Hughes suggests few points to make a test reliable. He lists 8 ways to make a test reliable.

- Adequate sampling: use more items. The more questions you ask, the more reliable the test becomes.
- Discrimination: Questions should be able to tell the difference between 'good' and 'weak' students.
- Restricted choice: don't give too many options.
- Clear items: Questions should be very clear free of ambiguity.
- Explicit instructions: Ensure students know exactly what to do.
- Familiar formats: Don't surprise students with a brand new layout. The question paper pattern should remain the same.
- Objective assessment: Use clear scoring rubrics.
- Multiple, independent scoring: have more than one person grade the work to avoid personal bias.

Relationship between reliability and validity:

Reliability is a pre-requisite for validity. A test cannot be valid if it is not reliable. Reliability is not sufficient. Just because a test is consistent does not mean it's measuring the right thing.

Authenticity of testing: with the shift in communicative language testing, authenticity is an additional principle that is seen to be important for validity. Performance on language tests should indicate how well the target language perform on language use tasks. Students perceive authentic tasks as more relevant and useful for the kinds of tasks they will have to do. The tasks designed to test should be real-

life activities to project relevancy among students. *A test is authentic if it reproduces a situation where a student must use the language to solve a real-life problem. E*.g. ordering food in a restaurant rather than just filling in verb tenses.

Interactiveness: It focuses on the test-takers engagement. It is the extent to which the test activates the student's internal mental and linguistic resources. Hight interactiveness means the student must use their language knowledge, background knowledge, and strategic mental processes to complete the task.

Practicality: It is the reality check of testing. It refers to whether the test is feasible to carry out given the resources available. A test may be perfect in theory, but if its too expensive, time-consuming, it is impractical. To identify you must evaluate.

Time: is there enough time to administer and crucially score the test?

Cost: Does the test stay within budgetary limits?

Human resources: Are there enough trained people available to proctor and grade the assessment. Supervise the exam etc.

Material Resources: are the necessary facilities available?

Ease of Administration: Are the directions clear, and is the scoring process straightforward?

Impact on teachers: Testing significantly shapes how teachers approach their curriculum and classroom methodology. It helps them in selection of curriculum, deciding teaching methods, teaching materials, designing activities and task. High stakes exams can put immense pressure on teachers to ensure students succeed, how well students do in the exams is a feedback to their teaching capabilities eventually.

Washback effect: The washback effect can deeply affect a student's psychological state and learning habits. Motivation and stress. Expectations of results may make students anxious. Some may get motivated.

Block I Unit Four: Techniques of Testing:

When we decide upon a technique for testing, we select the appropriate format for the technique. The technique and the format are guided by the specific purposes that we have for testing.

The main test types we will discuss are:

- Multiple choice
- True/false
- Matching the pair
- Gap-filling
- Cloze
- Error correction
- Short answer
- Long answer.

Multiple Choice items. (MCQ's) Making good MCQ is very challenging. Multiple Choice items are used to test reading and listening comprehension, grammar and vocabulary. It can also be used for testing writing and speaking but indirectly. MCQ technique is mainly used for recognition of knowledge. MCQ can have the following formats based on the language ability (Vocabulary, grammar LSRW etc.)

How Vocabulary can be tested using MCQ items:

- Isolated items (synonyms, antonyms, hyponyms)
- Items in a sentence context (synonyms)
- Items with two options in sentence contexts (words often confused)
- Items in paragraph contexts (appropriate words)

When words given to match or select in MCQ they test the cognitive level of Remembering and Understanding

Isolated items (synonyms, antonyms, hyponyms): remember in mcq there is one right answer and one wrong answer other two are distractors.

Example. (Isolated items in MCQ are context-free:
Tick the word that is closest in meaning to the given word.
Authentic:
a. Genuine b. honest c. false d. wise

Tick the word that does not belong to the same category as the given word. (on hyponym)

Vehicle:

a. Train b. car c. engine d. bus

Items in Sentence contexts: Context-embedded: the word is embedded in the context. The context provides clue to the meaning.

Tick the word that can replace the underlined word in the sentence so that the meaning remains the same choosing from the options given below

He presented the argument with great <u>lucidity.</u>
a. Transparency b. confusion. c. clarity d. emphasis

Choosing from two options in a paragraph:

Complete the sentence choosing the most appropriate of the two words in the given context.
The bank _____(refused, denied) to grant us a loan because we had no security.

Choice of words to fit the meaning contexts in paragraphs: In this item the examinee has to make sense of the whole paragraph and choose the most appropriate word to suit the context and bring about the intended meaning.

Complete the following paragraph choosing the most suitable of the four options given in brackets after each blank.

On a banana plantation snake were____(looked, held, felt, placed) in high regard…….

There are other items in vocabulary such as meanings of phrasal verbs, collocations, word families, root word, affixation etc that can be tested through multiple choice formats.

How sentence structure(grammar) can be tested using MCQ: Multiple choice questions are used for error identification. It tests both grammar and writing ability in an integrative manner. In standard test four parts are underlined and the part which has error has to be circled. E.g.

Four parts in the following sentence has been in underlined and numbered.
I abree that(a) a knowledge of Latin is helpful to build (b) a good English vocabulary, but (c) I don't think I have the capacity to study the subject at the moment. No error (e)

This is based on the assumption that a good language user will be able to identify errors in language, the implication being that he/she knows the right form.

How reading skills can be tested through MCQ items: the strategies we use for these levels of comprehension are skimming for overall comprehension, scanning for facts, intensive reading for inferential comprehension and word attack strategies for lexical comprehension.

The basic format consists of an incomplete statement followed by four options, one of which is to be selected to complete the statement.

Factual comprehension: The questions here are quite straightforward. The examinee has to read the text and get at the facts directly. Only one option completes the stem according to the meaning conveyed in the text. This is realized through Scanning.

On hot summer a shopkeeper was working in his shop. Business had been slow and there were no…

Tick the answer that best completes each of the following statements.

1. *When the woman entered the shop, the shopkeeper was*
   Looking at the goods in his shop
   Attending to his customers
   Examining his account books
   Looking at the woman outside the shop.

These types of questions where the reader is has to read the text and get the facts directly from text they test Factual comprehension of the reader which is achieved through scanning. There has to be structural similarity and uniformity of length of all the four options in each item.

Inferential Comprehension through MCQ: Although the format is same, the level of comprehension is different here.

Read the paragraph and complete the statement choosing one of the options given below.
When the District Commissioner arrived at Okonowo's compound at the head of…

When Obeirika repeated his answer 'He is not here'

a. *He showed that he was very simple and honest man.*
b. *He showed his defiance against the commissioner.*
c. *He still had not understood the commissioners' question*
d. *He meant more than the Commissioner could understand at*

The right meaning lies beyond the words and phrases used, in the content and the context of the story from which it has to be inferred.

Overall Comprehension: Global comprehension it is achieved through skimming. The MCQ can also be used for testing global meaning.

Read the poem 'The Road Not Taken'.
The poem is about.

a. Admiring nature
b. Making choices
c. Exploring the woods
d. Diverging forest roads.

Care should be taken to have the same structure in all the four options.

Lexical Comprehension: through MCQ.

Read the passage and choose the right meaning of the given word from the four options below each word.
The lorry had been ordered to arrive...
1. *Apparent*
   a. Obvious b. true c. strange d. well-known.

How listening skills can be tested through MCQ:

Sound discrimination, global comprehension, factual comprehension and evaluative listening can be tested through MCQ.

Sound Discrimination: Isolated items. We can limit this to two options.

A list of words will be read out to you. Tick the word you hear in each of the pairs you have on the printed sheet.

Pairs like 'later/latter' quite/quiet' can be given.

Global comprehension (listening) through MCQ. Listening to a text and understanding its gist; taking down the main points.

Factual comprehension (listening): Factual details dates, time, people's names, objects etc can be tested through listening skills through MCQs

Evaluative listening: Items involving listener's opinions on the speaker's intentions can be given.

Multiple Choice questions are also used for error identification as a part of writing tests. This is essential testing grammatical accuracy.

Coherence in writing can be tested by giving four sets of the sequence of sentences that form a paragraph. E.g.

Sentences: Put them in proper order.

A. As a result, many cities are now investing in public transportation.
B. Traffic congestion has become a serious problem in urban areas.
C. This move helps reduce pollution and saves commuters time.
D. The increasing number of private vehicles is the main cause of this issue.

Options:

a) B – D – A – C

b) D – B – A – C

c) B – A – D – C

d) A – B – D – C

Some of the features of MCQ testing technique: It is an indirect-testing method.

It tests only recognition of knowledge; no production of language is involved. It caters to the cognitive levels of 'remembering,' and 'understanding' Writing and speaking abilities can only be indirectly tested through this format. It allows guesswork and cheating may take place during exams. Does not test language ability. Not all areas of language lend themselves to multiple choice questions. Scoring is easy. It does not produce effective learning. Rote learning happens. Some points to remember about the structure of the options.

- They should have the same clause structure (all should verbs, or nouns)
- They should be of approximately the same length.
- If there is a stem to be completed, all the options should complete it.

Binary Choice: yes/no, True/false. It could also be Agree/Disagree, Acceptable/unacceptable. This is also a kind of multiple choice except that there only two options. The items are in the form of declarative sentences. 50/50 chances of getting the right answer therefore loses reliability. To make it challenging the learners are asked to justify their answer why yes, why no. etc.

Binary choice can be used for grammatical awareness of students asking the acceptability or unacceptability of the sentences. E.g.

State whether the following statement are Acceptable (A) or Unacceptable (U) according to Standard English.

- *The home minister says he is not averse to talk to the Maoist.*
- *The cause of Tsunami are not known.*

True or False format is mostly used for testing reading comprehension.

Given below is a text. Examine the two sets of true/false items A, and B given below based on the text.

You see a string of coolies coming along, one after another, each with a pole on his shoulders from the ends of which hand …PARAGRAPH.

State whether the following are true or false according to the text:

a. *Each coolie carrier a heavy pole with two bales hanging from its ends.*
b. *On their way, the coolies relax, smoking and chatting …*

For factual comprehension and lexical comprehension only, the factual information is required which can be accessed through scanning. For inferential comprehension the reader is required to multiple readings for deeper understanding. It asks the reader to share their opinions and beliefs; they must pass their judgement on the happening on the story within the paragraph.

Matching: Matching test item is easier because the number of options is almost equal to the number of the given items. One or two extra options are usually given to increase the challenge level. Matching items are very useful to testing vocabulary. E.g.

Match the proverbs in Column A with their meanings in Column B.

A	B
Make hay while the sun shines	A small sign is no indication that something is going to happen.
A stitch in time saves nine	It's better to have something which you already have than risk losing it trying to get something more.
A bird in hand is worth two in the bush	Prompt action prevents further problems.

Synonyms, antonyms and phrase meanings are the most common content areas that are tested through this format. It is also used for testing reading comprehension at the early reading stages. Also for sentence structure. When selecting items for synonyms or antonyms, take care to see that the word endings do not give away clues to the right answer. Have all the items in the same part of speech.

Match the words in Column A with their antonyms in Column B.

A	B
Notorious	Thrifty
Generous	Impolite
Extravagant	Flexible
Courteous	Famous
Stubborn	Stingy

In both the columns the words given are adjectives which increases the challenge for students. When giving matching items in antonyms, the

answers should not be ambiguous. Each item should have the exact matching equivalent. The most important thing for match the pairs is that in both columns the words should be of same part of speech. The main drawback of the matching format it that it is mostly used to tests construction in isolation. It mainly tests recognition knowledge -it caters to the cognitive levels of 'remembering and understanding.'

Gap-Filling: this is a kind of limited production technique. It is good for factual reading and listening comprehension. The answers can be in one or two words or a phrase. There are various formats of gap-filling.

Sentence Completion with blanks at the end of the sentence. The blanks can be one or two indicating the number of words to be filled in.

Read the text given below and complete the statements that follow.

Rainforests are forests characterised by high rainfall. The minimum normal annual rainfall for a forest…

The minimum annual rainfall in a rainforest
_____
The monsoon trough is also known as
_____

This is useful for actual reading and listening comprehension.

Fill in the Blanks: It is an extension of gap-filling. Blanks can appear anywhere in a sentence or paragraph. Grammar items, pronouns, prepositions, articles and conjunctions are commonly tested using this format.

Fill in the blanks can also be very useful in testing listening skills.

Listen to the announcements and fill in the blanks with what you think was said.

1. *Train no. 2710 will arrive on Platform no.___*
2. *We are sorry to announce a delay in the departure of Train No. 2710 to .. The train leaves at ____*

This kind of listening for factual information is realized through scanning.

Another important aspect of listening is distinguishing between similar sounds. Fill in the blanks can be used for this. Choice between two options can also be provided. Here is a listening text with blanks to be filled based on it.

He was a most wonderful sight. Though a little hole in the ____ children had crept in..

We can have an alternative form. On Discriminating sound on listening skills.

1. He was a most wonderful ___(light/sight) 2. Through a little hole in the _____(hall/wall)

Filling in the blanks format is commonly used for grammar and vocabulary. This would test knowledge of forms. It involves limited production based on knowledge of forms.

Cloze: Cloze is a form of gap-filling. It tests integration of skills for language use. In the typical format, every N^{th} word is omitted; that is every third or fourth or fifth word is omitted and the blanks have to be filled I n with words appropriate to the context.

In the below cloze every fifth word is omitted.

A solar eclipse occurs__ the moon comes between_____ sun and the earth____

The fifth word has to be predicted. The ability to predict what the sequence of words could be is a part of an individual's language awareness. It is valid for overall language ability.

C-test is another variation of the cloze. Like the cloze it is an integrative testing instrument that tests overall language competence. In this technique the second half of every second word is omitted. E.g.

Any move___ draws t___ eyes o____ the audi____

In a C-test, it is usual to give the first and last lines completely in order to help the reader understand the context. For both the reader has to read the whole text and understand the overall import in order to successfully arrive at the word completion.

Error correction: error is usually discussed under multiple choice for sentence structure. In standardized tests errors have to be identified and the correct form rewritten in the blank provided against the line in which the error occurs.

An example follows:

I can never forget one of my childhood experience _____ I was trying to climb a tree when my foot slip and I _____

In this format, the error has to be identified and rectified. Recognition and production are tested. The errors are mainly grammatical.

Short Answer: Traditionally, reading comprehension has been tested through short Wh-questions. **Short answers are open-ended hence scoring will necessarily have an element of subjectivity.**

Read the text below and comment on the two sets of questions below it.

Read the following text and answer the questions below.

Governments have seized on tourism as a way of creating employment…

Answer the following questions in one or two sentences:

Set A:

1. Why have governments seized on tourism?
2. What do tourists spend money on?
3. Whose money circulates in the economy?

These kinds of questions require factual reading through scanning. The answers can be derived without with a literal understanding of the text.

Set B:

1. What are the economic benefits of tourism?
2. How does tourism help in bringing in foreign exchange?
3. What increases the value of a country's culture for the tourist?

These questions demand thinking and reformulating. The answers require global comprehension which can be realized through skimming. This set of questions require the testees to locate the answer and present it differently.

If we are interested in testing writing ability using a reading passage, then the input for the writing is drawn form the reading content, and the responses reflect the measure of writing ability. In this case, we are not looking for the right answers but how well the testee is able to communicate through written language.

Long answer: Long answers test integration of skills and are a dependable indicator of overall language ability. Examples of long

answer test items are, essays, letters, summaries, etc. Questions based on the literature section also elicit long answers. In order to ensure validity of scoring long answers, criteria have to be defined accurately. The following criteria is followed for assessing long answers.

- Content (relevance and originality)
- Organization of thoughts and words
- Coherence
- Cohesion
- Grammatical accuracy (and spelling)
- Choice of vocabulary
- Overall presentation

- MCQ is used in standardized tests, it focuses on exactly on the construct sought and ensures quick and reliable scoring.
- Matching is a useful test format item used for testing vocabulary and sentence structure. Care is necessary to see that the forms in two column are uniform.
- Gap-filling is used in almost all the skill areas and the testing of grammar and vocabulary.
- Cloze is a test of overall language ability, similar to gap-filling but based on more a comprehensive understanding of how language functions in communication.
- While objective formats have their advantages, they are indirect measures of testing. Productive ability, integrating all the elements of language, is possible only through extended writing and speaking activities.